Test Planning for Observational Studies

Maj Brian Stone
HQ AFOTEC
Release Date: 12 Mar 18

Overview

- Background on Observational Studies
- Methodology for Test Planning
- Poisson Processes
- Examples

Background

- Observational Study – A study (test) conducted where the independent variables (factors) are not under the control of the researcher (test team)
- The primary reason AFOTEC uses observational studies is because the system under test (SUT) is already operational and factor levels cannot be deliberately changed
  - Changing levels may be infeasible
  - Changing levels may adversely affect the operational environment
- Observational studies are common for enterprise software systems:
  - Defense Enterprise Accounting and Management System (DEAMS)
  - Air Force Integrated Personnel and Pay System (AFIPPS)

Background

- Although factor levels cannot be controlled for the purposes of the test, a test matrix that defines the operational battlespace is still useful for test planning
  - The test design matrix is used to determine where and when to send the test observers to observe and gather data
  - The test team attempts to observe as many factor level combinations as possible during the test event
- This is a portion of the test design matrix for the Air Force Integrated Personnel and Pay System (AFIPPS)

Design Point	Business Process	User Role	Affected Member Component	Recipient Classification	Level of Input
1	BP00180.1 Record Disciplinary Actions	HR Specialist	Regular Air Force	Cadet	Base
2	BP00180.1 Record Disciplinary Actions	HR Specialist	Regular Air Force	Enlisted	Base
3	BP00180.1 Record Disciplinary Actions	HR Specialist	Regular Air Force	Officer	Base
4	BP00153.1 - Process Retirement	HR Specialist	Air Force Reserve	Enlisted	ARPC
5	BP00153.1 - Process Retirement	HR Specialist	Air Force Reserve	Officer	ARPC
6	BP00970.1 Manage Leave Requests	Approval Authority	Regular Air Force	Cadet	Unit
7	BP00970.1 Manage Leave Requests	Approval Authority	Regular Air Force	Enlisted	Unit
8	BP00970.1 Manage Leave Requests	Approval Authority	Regular Air Force	Officer	Unit

Problem Statement

- This test design matrix shows the factor labels across the top row
- Each column contains the levels of a factor
- The test will be conducted as an observational study, as any business practices conducted in AFIPPS will affect actual personnel records and human resource (HR) databases

Design Point	Business Process	User Role	Affected Member Component	Recipient Classification	Level of Input
1	BP00180.1 Record Disciplinary Actions	HR Specialist	Regular Air Force	Cadet	Base
2	BP00180.1 Record Disciplinary Actions	HR Specialist	Regular Air Force	Enlisted	Base
3	BP00180.1 Record Disciplinary Actions	HR Specialist	Regular Air Force	Officer	Base
4	BP00153.1 - Process Retirement	HR Specialist	Air Force Reserve	Enlisted	ARPC
5	BP00153.1 - Process Retirement	HR Specialist	Air Force Reserve	Officer	ARPC
6	BP00970.1 Manage Leave Requests	Approval Authority	Regular Air Force	Cadet	Unit
7	BP00970.1 Manage Leave Requests	Approval Authority	Regular Air Force	Enlisted	Unit
8	BP00970.1 Manage Leave Requests	Approval Authority	Regular Air Force	Officer	Unit

- Problem
  - Factor level combinations do not occur with equal frequency
  - The test team needs to know how long to schedule an observational test event in order to observe all factor level combinations

Methodology

- **For every observational test event, estimate the frequency of the factor level combinations**

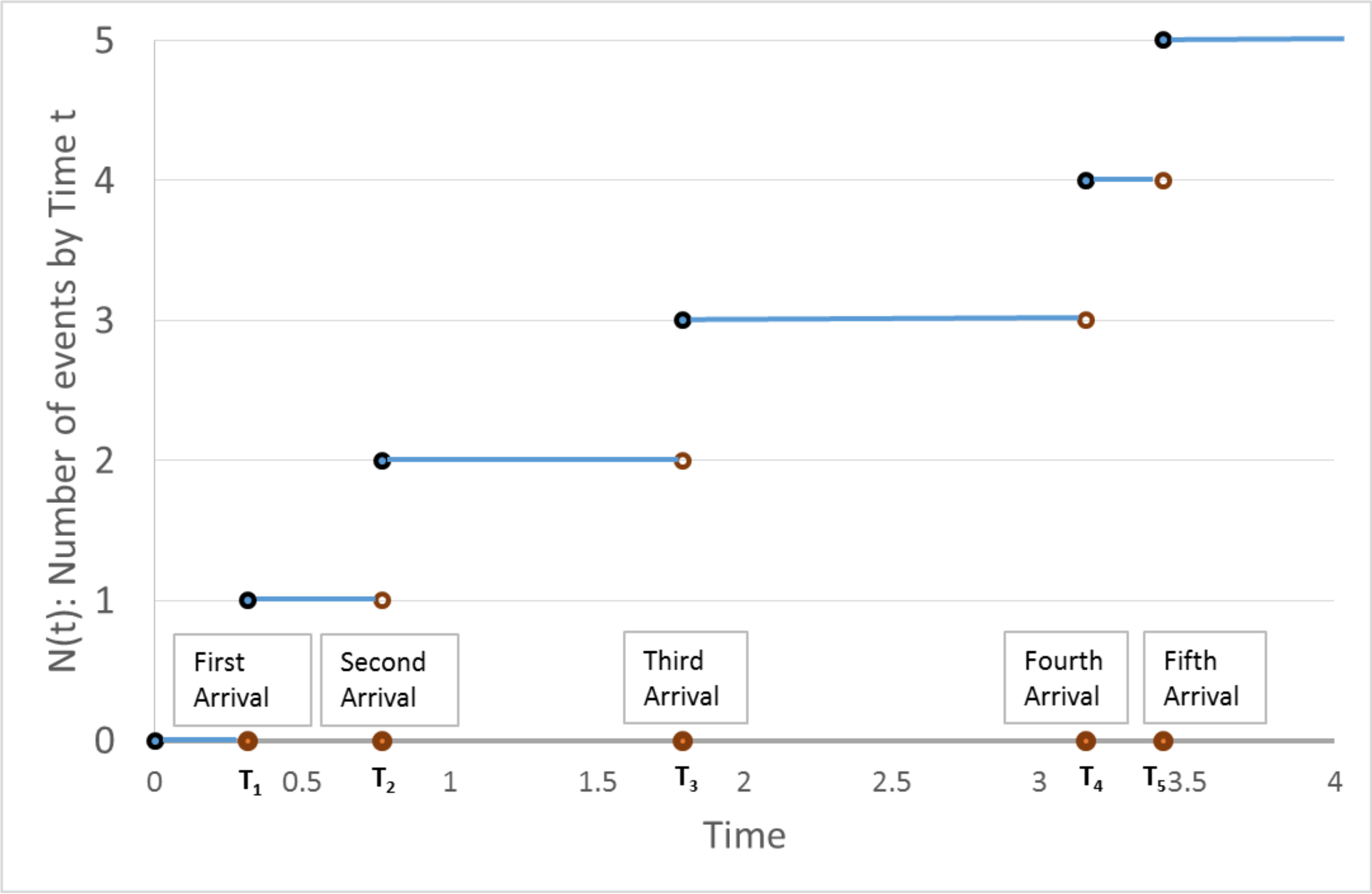
Design Point	Business Process	User Role	Affected Member Component	Recipient Classification	Level of Input	Frequency of Event
1	BP00180.1 Record Disciplinary Actions	HR Specialist	Regular Air Force	Cadet	Base	1 Per 160 Hrs
2	BP00180.1 Record Disciplinary Actions	HR Specialist	Regular Air Force	Enlisted	Base	3 Per 40 Hrs
3	BP00180.1 Record Disciplinary Actions	HR Specialist	Regular Air Force	Officer	Base	3 Per 40 Hrs
4	BP00153.1 - Process Retirement	HR Specialist	Air Force Reserve	Enlisted	ARPC	2 Per 40 Hrs
5	BP00153.1 - Process Retirement	HR Specialist	Air Force Reserve	Officer	ARPC	2 Per 40 Hrs
6	BP00970.1 Manage Leave Requests	Approval Authority	Regular Air Force	Cadet	Unit	3 Per 40 Hrs
7	BP00970.1 Manage Leave Requests	Approval Authority	Regular Air Force	Enlisted	Unit	150 Per 40 Hrs
8	BP00970.1 Manage Leave Requests	Approval Authority	Regular Air Force	Officer	Unit	100 Per 40 Hrs

- **Use a Poisson Process model to determine the time required to observe the rare events**

Siméon Poisson
1781-1840

The Poisson Process

- **The Poisson Process is a common model for scenarios where we are counting the occurrences of certain events that appear to happen at a certain rate, but completely at random**



The Poisson Process

- **Examples of scenarios modeled by Poisson Processes:**
  - **Arrivals of customers at a store**
  - **The number of car accidents in a certain area**
- **Poisson Processes have a rate parameter $\lambda > 0$ that describes the expected number of events to occur in some interval of time**
  - **The number of events in any interval of length *t* is a Poisson random variable with parameter (or mean) λt**
  - **Example: If 5 events happen per hour on average, then $\lambda = 5$, and the number of events occurring in *2* hours is a Poisson random variable with mean $= \lambda t = (5)(2) = 10$**
- **There are a few assumptions:**
  - **The events occur at a certain rate but completely at random within an interval**
  - **The number of events in some interval of length *t* is independent of the number of events in another (disjoint) interval of length *t***

The Poisson Process

- We can use a Poisson Process model to determine the probability that an event will occur by time *t*
- Let $N(t)$ be a Poisson Process with rate λ
- Let T_1 be the time of the first occurrence of a factor level combination we want to observe in an observational test
- Then $P(T_1 > t) = e^{-\lambda t}$, i.e. T_1~Exponential(λ)
- We can use this model to determine the time, *t*, required for a test, given that we want a *Y%* chance to see a rare factor level combination
- A Y% chance to see an event implies $P(T_1 > t) = 1 - \left(\frac{Y}{100}\right)$

$$t = \frac{-ln\left(1 - \left(\frac{Y}{100}\right)\right)}{\lambda}$$

Example 1

- **Suppose a factor level combination occurs randomly 1 time every 3 working days (on average) while a system is operating**
- **We want to determine how long to conduct an observational study so that we have an 80% chance of observing the event**
- **First determine λ (Assume an 8 hour working day)**

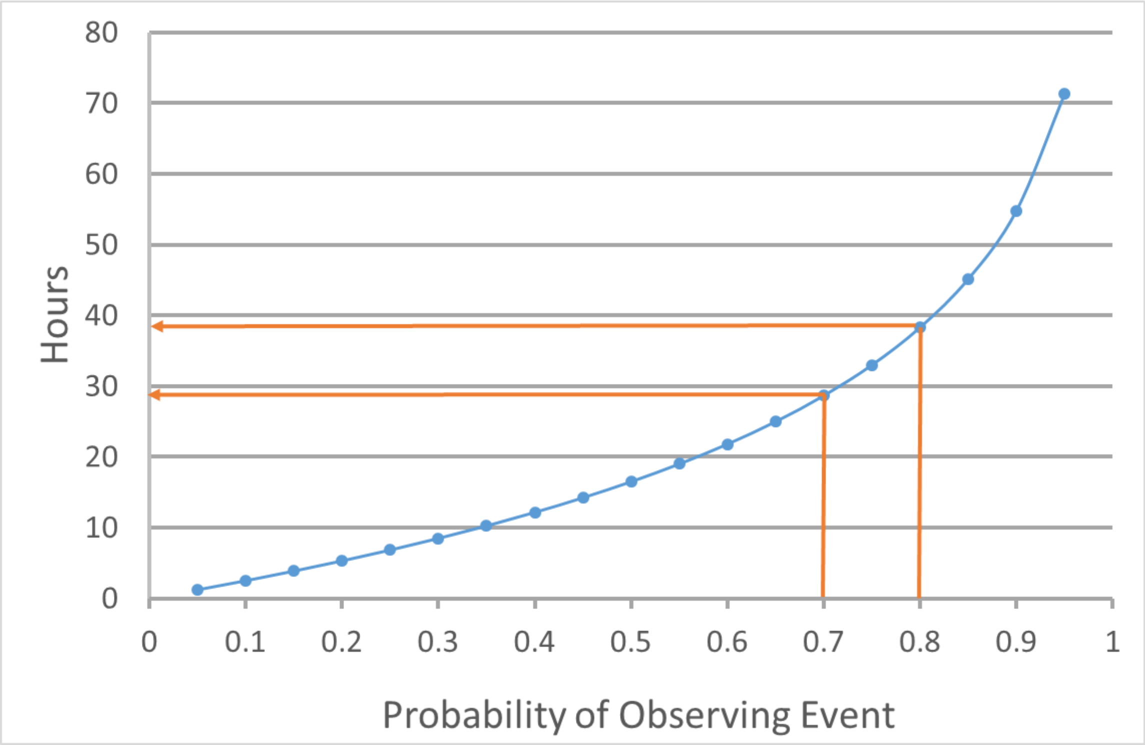
$$\frac{\mathbf{1}\ \boldsymbol{event}}{\mathbf{24}\ \boldsymbol{hours}} \Rightarrow \lambda = \mathbf{0.042}\ \boldsymbol{events/hour}$$

- **Next calculate *t*, the required test time**

$$\boldsymbol{t} = \frac{-\boldsymbol{ln}\left(\mathbf{1} - \left(\frac{\mathbf{80}}{\mathbf{100}}\right)\right)}{\mathbf{0.042}} = \mathbf{38.3}\ \boldsymbol{hours}$$

Example 1

- It is possible to do sensitivity analysis on the probability requirement
- If we only require a 70% probability of observing the event, we can cut nearly 10 hours off the test time



Example 2

- **Suppose we want to observe 3 occurrences of a factor level combination**
- **The interarrival times between events in a Poisson Process are Exponential(λ) random variables (RV); therefore the arrival time of the 3rd event (T_3) is the sum of 3 Exponential(λ) RVs**
- **Hence T_3 is distributed Erlang(3, λ)**
- **The probability that $T_3 > t$ is:**

$$P(T_3 > t) = \sum_{n=0}^{3-1} \frac{1}{n!} e^{-\lambda t} (\lambda t)^n$$

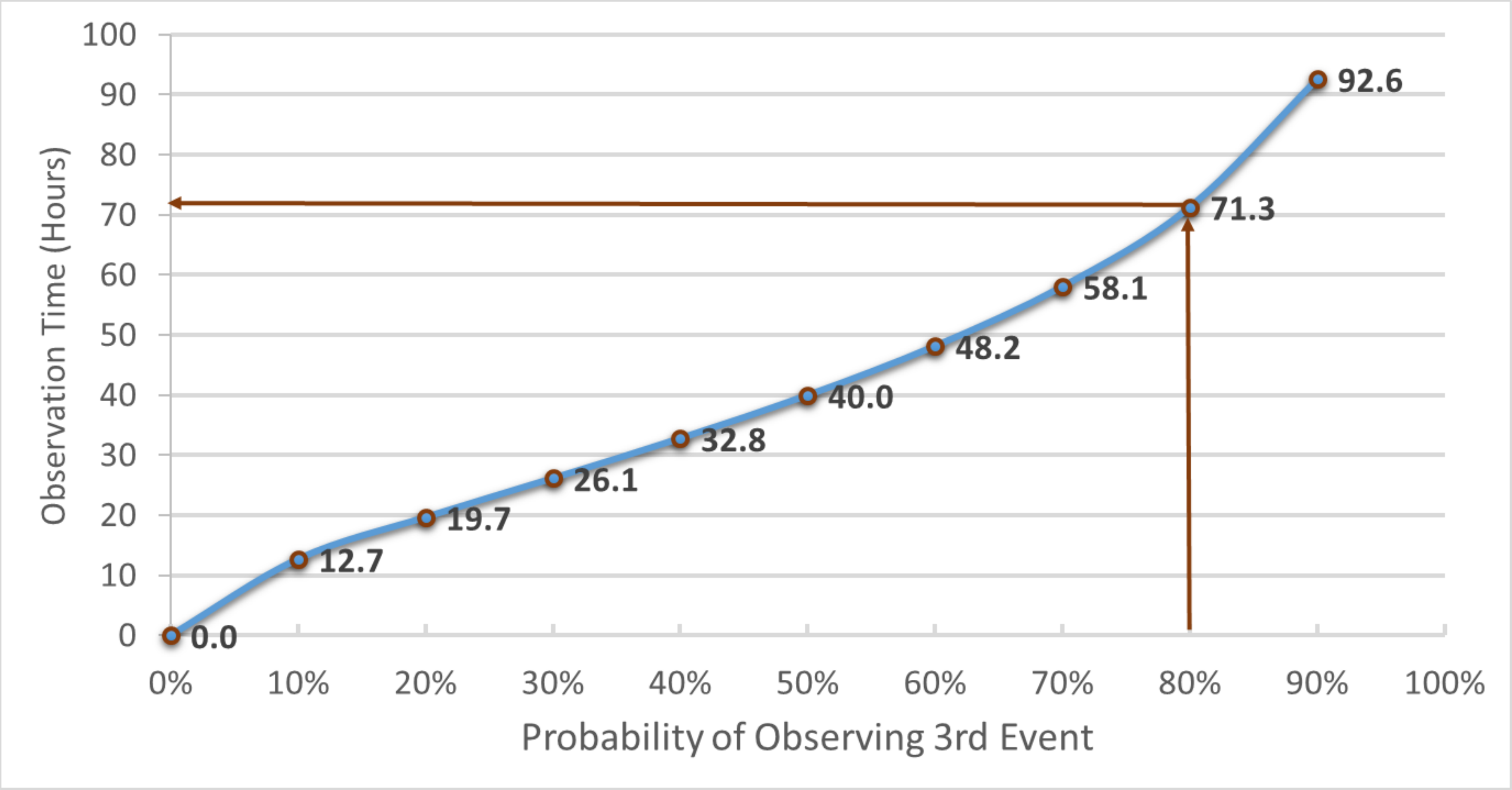
- **Therefore the probability that $T_3 > t$ when $\lambda = 0.042$ is:**

$$P(T_3 > t) = \frac{1}{0!} e^{-\lambda t} (\lambda t)^0 + \frac{1}{1!} e^{-\lambda t} (\lambda t)^1 + \frac{1}{2!} e^{-\lambda t} (\lambda t)^2$$

$$P(T_3 > t) = e^{-0.042t} + e^{-0.042t}(0.042t) + \frac{1}{2} e^{-0.042t}(0.042t)^2$$

Example 2

- The observation time (in hours) required to observe 3 events was calculated for various probabilities (0%, 10%, 20%,...,90%)
- 71.3 hours are required for an 80% chance of observing 3 events



Example 3

- Suppose we have 10 working days to conduct our observational study; The Poisson Process model can be used to determine the probability of seeing a rare factor level combination

Design Point	Business Process	User Role	Affected Member Component	Recipient Classification	Level of Input	Frequency of Event
1	BP00180.1 Record Disciplinary Actions	HR Specialist	Regular Air Force	Cadet	Base	1 Per 160 Hrs
2	BP00180.1 Record Disciplinary Actions	HR Specialist	Regular Air Force	Enlisted	Base	3 Per 40 Hrs
3	BP00180.1 Record Disciplinary Actions	HR Specialist	Regular Air Force	Officer	Base	3 Per 40 Hrs

- There are 80 hours in the test; Now $P(T_1 < t) = 1 - e^{-\lambda t}$

$$\frac{1\ event}{160\ hours} \Rightarrow \lambda = 0.00625\ events/hour$$

$$P(T_1 < 80) = 1 - e^{-(0.00625)(80)} = 1 - 0.607 = 0.393$$

- There is a 39.3% chance of observing the rare event

Summary

- Formula to calculate the test time *t* required to have a *Y%* chance of seeing a factor level combination:

$$t = \frac{-ln\left(1 - \left(\frac{Y}{100}\right)\right)}{\lambda}$$

- Formula to calculate the test time *t* required to have a *Y%* chance of seeing *k* occurrences of a factor level combination:

$$1 - \left(\frac{Y}{100}\right) = \sum_{n=0}^{k-1} \frac{1}{n!} e^{-\lambda t} (\lambda t)^n$$

- Formula to calculate the probability of observing an event by time *t*:

$$P(T_1 < t) = 1 - e^{-\lambda t}$$

Questions